A Bayesian Approach to Causality Assessment 1: Foundations

by
David A. Lane, Ph.D., Tom A. Hutchinson, M.B.,
Judith K. Jones, M.D., Michael S. Kramer, M.D.,
and Claudio A. Naranjo, M.D.[1]
University of Minnesota

Technical Report No. 461
February 1986

Direct correspondence to: Professor David A. Lane, School of Statistics, 270 Vincent Hall, University of Minnesota, Minneapolis, MN 55455 (612-373-3035)

---

[1] This research has been supported by grants from the University of Minnesota Graduate School, the Drug Information Association, the American Medical Association Education and Research Foundation and Ciba-Geigy Canada. Exept for the principal author, the authors are listed alphabetically. The principal author wishes to acknowledge the helpful comments and criticisms of his colleagues Tom Rector and Bill Sudderth.

Abstract

A new appproach to the problem of assessing causality for adverse drug reactions is presented. The approach is based on Bayesian probability theory, and it is designed to anwer the following question in a logically satisfying and nonarbitrary way: given all the available information, what is the probability that a given adverse clinical event was caused by some particular drug to which the patient had been exposed? The approach is illustrated by a case in which amoxicillin is suspected of causing diarrhea. Although much work remains to be done before the approach can be easily implemented, it is argued that the approach satisfies basic criteria for causality assessment methods, a claim that cannot be made for any other currently available technique.

Key words: adverse drug reactions, causality assessment, probability, coherence

A BAYESIAN APPROACH TO CAUSALITY ASSESSMENT I: FOUNDATIONS

The answers to important clinical, research and policy questions can depend in part on the extent to which one is justified in believing that particular adverse clinical events were caused by specific drugs. For example, should a clinician discontinue the use of an effective antiinflammatory drug in a patient to whom it may be causing anginal pain? How should a pharmaceutical manufacturer react to the fact that two patients in a clinical trial of a new antiulcer drug developed liver disease? Should a trial of a new heart failure drug be terminated because four of nine patients on the drug died suddenly within the first three months of treatment? How should an epidemiologist studying the incidence of in-hospital iatrogenic diseases decide whether a particular case of renal failure is drug-induced?

All these questions involve the causality assessment problem: given all the available information, what is the probability that a given adverse clinical event was caused by some particular drug to which the patient had been exposed? In a previous paper [14], we developed criteria that methods for solving the causality assessment problem should satisfy, and we argued that no current method satisfied these criteria. In this paper, we describe a new approach to the causality assessment problem, and we show that this approach satisfies the criteria discussed in [14]. Ideas related to the new approach can be found in [13] and [1].

The new approach is based on Bayesian probability theory. In section 1, we explain the essential ideas of this theory and present reasons why it should be applied in the causality assessment context. The new approach is developed in sections 2 and 3; in section 4 it is applied to a case in which amoxicillin is suspected of causing diarrhea; and in section 5, it is discussed in relation to the criteria of [14].

1. Probability, Coherence and Causality Assessment

We believe that two features of the causality assessment problem contribute substantially to its difficulty. The first of these is uncertainty. The assessor is usually uncertain about many of the key elements he must integrate into his assessment. His uncertainty may be about some of the facts of the case (did the patient actually take drug D before event E began? has the patient ever experienced an event similar to E before?); about background information that affects how these facts are to be interpreted (how long should it take before toxic quantities of the relevant metabolite of D have accumulated at the target organ? what is the incidence of events like E in patients similar to the present one, but who have not taken D?); or about the assumptions that it is appropriate to make when determining the evidentiary significance of the clinical data (if E is an adverse reaction to D, is the mechanism dose-dependent or immunologic? could E be a clinical sequela of the disease for which the patient is taking D as treatment?). Somehow, the assessor must take into account this uncertain information and the extent of his uncertainty when he evaluates the probability that D caused E.

The second feature has to do with the complexity of the information that affects causality assessment. There is always more than one stream of evidence that must be merged in any causality assessment problem. Several different factors are relevant to any causality

assessment problem, and the evidence about each of them usually comes from several different sources. The factors include background incidence of events like E (in patients who had previously taken D, as well as in those who had not), aspects of the patient's history (including previous experience with drugs similar to D and events similar to E, as well as demographic and genetic information), timing of the event in relation to administration of D, clinical characteristics of the event (including drug levels in tissues or body fluids), and the patient's response to dechallenge (withdrawal of D) or rechallenge (readministration of D), when these occur. Information sources include the assessor's previous clinical experience and his clinical judgement, clinical observations and laboratory findings on the patient in question, data from epidemiologic studies, as well as facts and theories from pharmacology and other basic sciences. Frequently, information from one source or about one factor conflicts with information from another source or about another factor. Even when the information is not mutually contradictory, some way must be found to weigh the significance of each of the pieces and combine these "weights of evidence" in a reasonable way.

Both of these problems can be addressed with the help of Bayesian probability theory. This theory provides a set of normative rules for reasoning in the face of uncertainty. The rules are not arbitrary, in the following sense: violate them, and you act like someone who simply throws money away, with no gain to himself. On the other hand, while following the rules guarantees that one will act consistently with his own opinions about matters relevant to the problem at hand, it does not ensure that these opinions are correct. Fortunately, treating the causality assessment problem from a Bayesian point of view has an additional benefit: the method provides a unique and logical connection between the overall causality assessment and a set of other judgements that are subject, at least in principle, to empirical checks, so that the assessor can (again, in principle and with patience) discover if he is wise as well as consistent. The rest of this section elaborates on the claims presented in this paragraph.

We begin with a definition of "probability." As we explained in [14], we regard probability as a subjective measure of degree of belief. In causality assessment, as in many other real-world problems, the ultimate goal is to decide how to act when faced with uncertainty. Thus, the aspect of belief that we would like probability to measure is the subject's propensity to act as though the proposition whose probability he is evaluating is true. One way to achieve this, following the pioneering work of the Italian mathematician Bruno de Finetti [4,5], is to reduce the assessment of uncertainty to an economic decision, where the acts to be taken and the values of their consequences are clear. To this end, evaluate the probability for you of a proposition A according to the following thought experiment: decide upon a number p such that you are neutral between buying and selling for $p a ticket that will be worth $1 if A is true and otherwise it will be worth nothing; this number p is your probability that A is true (you must imagine that at some specified time in the future, you will find out for sure whether A is or is not true).

Note that if you are sure that A is true, p must be 1, and if you are sure that A is false, p must be zero; otherwise, there is a unique number p between 0 and 1 satisfying the definition (for any number greater than p, the price is too high, and you would be unwilling to buy the ticket; while you would not agree to sell the ticket for less than p). Thus, although

it may be difficult to determine p exactly, just as it is difficult to decide exactly how much you would be willing to pay for a new house, in principle both quantities exist, and, at the least, bounds on both could be determined by how you act when betting or negotiating for the house.

Some people find it easier to think about odds than probability. Giving 3 to 1 odds in favor of A is exactly the same as setting its probability at 3/4 ; generally, a probability of p is equivalent to (favoring) odds of p/(1 - p).

With de Finetti's definition of probability, it is possible to give a precise meaning to inconsistent reasoning in the face of uncertainty. Suppose you simultaneously assess the uncertainty you feel about many different propositions that are related in various ways. Using de Finetti's definition as your measure of uncertainty, you have therefore simultaneously set the price for many tickets. Is it possible, in principle, that someone could transact with you for some of these, at your prices, in such a way that you must pay out more than you receive from him, no matter which of the propositions are true and which false?[1] If so, in your assessments you have in effect made economic decisions with unacceptable economic consequence, certain financial loss. The possibility of such loss is a concretization of the inconsistent reasoning that underlies it.

A set of bets that makes money no matter what happens is called a Dutch book in gambling circles. The rules of Bayesian probability theory guarantee that all your probability assessments fit together in such a way that a Dutch book cannot be made against you. In this sense, you either reason about uncertainty consistently with these rules -- or you act like someone who is willing to give money away without any chance of getting it back. And, since your probabilities for the various propositions exist whether you determine them or not, this result remains true whether you explicitly and quantitatively assess your uncertainty or you just do it implicitly and qualitatively, as in most current approaches to causality assessment. A set of probability assessments that is consistent with the rules is called coherent; any reasoning about uncertainty that is not consistent with any coherent set of probability assessments is called incoherent.

To describe the rules of Bayesian probability theory, we need to introduce one more concept, conditional probability. To understand the idea of conditioning, suppose you are considering whether or not to purchase a house that you have just hired a structural engineer to examine. The price you would be willing to pay for the house will depend on whether the

---

[1] As an example, suppose A represents the proposition that it will rain tomorrow, and A^c that it will not rain tomorrow. There is nothing inherently wrong with assessing $P(A) = 0.25$, nor with assessing $P(A^c) = 0.25$ (though either one taken alone could well be unwise -- say, in a desert or a rain-forest). On the other hand, it is clearly inconsistent to assess both quantities as 0.25: if you did so, and sold tickets at the assessed prices, someone could buy one of each ticket for a total outlay of 50¢ -- and give them back to you today demanding $1 in return, since one or the other proposition is bound to be true tomorrow, and hence one of the tickets must be worth $1, while the other will be worth nothing. The 50¢ sure loss you, the assessor, face in this situation reveals the inconsistency in the simultaneous assessments of 0.25 for the two probabilities $P(A)$ and $P(A^c)$.

engineer's report is favorable or not. Thus, to decide how to proceed in your negotiations, you could now determine a price you would be willing to pay assuming the engineer produces a favorable report and a price assuming an unfavorable report. These are conditional prices. They become operative when a specified condition is found to be true. If the condition turns out to be false, the price determined conditionally on its truth becomes irrelevant.

The same idea applies in the probability assessment context. Informally, the conditional probability of a proposition A given a proposition B, denoted P(A|B), is the degree of belief that the assessor has in A if he agrees to assume that B is true (that is, if he acts as though he is certain about B). Formally, we can define P(A|B) as the price of a ticket worth $1 if both A and B are true, worth nothing if B is true and A is false, and the purchase price is refunded and the bet called off if B turns out to be false (that is, the bet is made conditionally on the truth of B). All probabilities are really conditional probabilities, since the assessor always conditions on everything that he believes to be true with certainty (that is, the unconditional probability P(A) is the same as P(A|B), where "B" refers to the facts and opinions that the assessor regards as true and on the basis of which he determines his uncertainty about A).

Sometimes, we consider probabilites that are conditional on more than one proposition. For example, "P(A|B and C)" or equivalently "P(A|B,C)" denotes the probability of A given B and C, and it represents the price for a bet on the truth of A that is called off unless both B and C turn out to be true. When reading probability notations, remember that everything inside the parentheses to the left of the "|" is part of the proposition about which the assessor is measuring his uncertainty, and everything to the right of the "|" represents a proposition that, for the purposes of that particular assessment, the assessor is regarding as certainly true.

Now we can state the main result of Bayesian probability theory. Denote by "B" everything that the assessor regards as true with certainty. Then a collection of probability assessments is coherent if and only if the following three conditions hold:

1) The Normalization Condition: For every proposition A , P(A|B) must be between 0 and 1, inclusive.

2) The Additivity Condition: If A and C are mutually contradictory propositions,

$$P(A \text{ or } C \mid B) = P(A|B) + P(C|B).$$

3) The Multiplicative Condition: For any propositions A and C,

$$\begin{aligned} P(A \text{ and } C \mid B) &= P(A|B) \times P(C|A \text{ and } B) \\ &= P(C|B) \times P(A|B \text{ and } C). \end{aligned}$$

Here are two important consequences of these conditions that we shall use repeatedly in what follows:

1) Bayes' Theorem: For any propositions A and C such that P(C|B) > 0,

$$P(A|B \text{ and } C) = \frac{P(A|B) \times P(C|A \text{ and } B)}{P(C|B)},$$

or equivalently,

$$\frac{P(A|B \text{ and } C)}{P(A^c|B \text{ and } C)} = \frac{P(A|B)}{P(A^c|B)} \times \frac{P(C|A \text{ and } B)}{P(C|A^c \text{ and } B)},$$

where A^c is the proposition that states that A is false. The second expression of Bayes' Theorem is in odds form: the left-hand side is called the posterior odds in favor of A (that is, posterior to C), the first term on the right-hand side is called the prior odds in favor of A (that is, prior to C), and the second term on the right-hand side is called the likelihood ratio.

2) The Law of Total Probability: Suppose $A_1, A_2, \ldots, A_n$ are mutually contradictory propositions such that one of them must be true. Then for any proposition C,

$$P(C|B) = [P(C|B \text{ and } A_1) \times P(A_1|B)] + \ldots + [P(C|B \text{ and } A_n) \times P(A_n|B)].$$

Bayes' Theorem follows from the second equality in the multiplicative condition, and the Law of Total Probability follows from the additive and multiplicative conditions taken together. See appendix 1 for a detailed derivation of these two results.

How do these results help us address the causality assessment problem? First, the causality assessment problem itself is just a special kind of conditional probability evaluation. Given all available case and background information, one is required to assess his probability for the proposition that the drug D caused the event E (which we shall hereafter denote as "D->E"). To evaluate this probability, one must take into account all the data that matter, as well as what it is that makes them significant. According to the Bayesian theory, to reason coherently about all these things, one should first decompose the "global" causality assessment problem into a series of component problems, each of which deals only with information about one factor or source of information. For each of these simpler problems, one must measure his uncertainty about the facts and their evidentiary significance using de Finetti's definition of probability. Finally, the theory shows that there is only one way to piece together the solutions to these component problems into a coherent overall causality assessment, and that is to use Bayes' Theorem and the Law of Total Probability. Thus, the theory prescribes how we should deal with uncertainty and how we must merge different streams of evidence, and it points up the penalty should we fail to follow its prescription: we will necessarily reason, and act, incoherently.

There is still one important issue to address: how can you know if a probability appraisal is "right"? If probability measures subjective degree of belief, then this question seems meaningless (unless it addresses the problem of how someone else could determine whether the assessor is honestly announcing his real opinions, a problem that we shall not take up here). But this is not the end of the story, since probabilities of propositions about future observables can be converted into predictions about the values that these observables will assume, and the validity of the predictions can be subjected to empirical checks. For example, suppose two different coherent probability assessment methods are used to produce probabilities for a set of propositions about future observables, and one method consistently assigns higher probabilities to the propositions that turn out to be true, and lower probabilities to the ones that turn out false, than the other method. Then it seems reasonable to conclude that even though both methods are coherent, one of the methods is more in tune with the world than the other, and in the future (everything else remaining reasonably the same), one would want to modify his own opinions to concur with probabilities generated by that method rather than by its less effective alternative.

Now the proposition "D->E", whose probability it is the business of causality assessment to evaluate, is certainly not about any future observable. It is retrodictive rather than predictive in character, and typically whether D actually caused E in that particular case or not will never be known with certainty. Thus, the validity of a causality assessment method is not subject to a direct empirical check, in which the probabilities it produces are converted into predictions and the accuracy of these predictions is determined. Nonetheless, using Bayes' Theorem, it is possible to convert the causality assessment problem into a series of probability assessments, the propositions in each of which are about the values of future observables. This conversion will be achieved in the third section of this paper. Its implications are extremely important: in principle, the Bayesian approach to causality assessment allows the logical incorporation of a series of methods for evaluating components of the overall uncertainty about drug causation, and each of the methods can be subjected to empirical tests of its soundness.[2]

We conclude this section with a discussion of a frequently-raised objection to the use of probability theory for problems like causality assessment. Many people believe that it is impossible to assign a definite number to their uncertainty about the truth of a proposition. When asked, for example, what is the chance that a patient with a certain clinical condition treated in a particular way will experience some particular untoward clinical event in some specified time period, they respond that they just do not know; to attach a number to their uncertainty would be to introduce a meaningless and false precision to something essentially vague and even unknowable. We believe that this position is incorrect, for the following three reasons:

1) Using de Finetti's definition of probability, the precision in measuring probability is not false. There is certainly some justice in applying the charge of false precision to some

[2] We hasten to point out that these methods have yet to be developed; to the extent that the Bayesian approach can prove its value in applications, we are confident that they will be.

previous quantitative approaches to causality assessment, that have used uncalibrated, analogue probability-like scales ranging from 0 to 1 (or 100) [12,19], for which the resulting numbers have no clear interpretation and hence no real meaning (as discussed in [19]). But this is not true for the ideas developed in this paper, since the de Finetti definition gives a precise meaning to the measurement of uncertainty. Furthermore, Bayesian probability theory shows how optimal decision-making in the face of uncertainty depends precisely on probabilities defined in this way [15]. In addition, any optimal decision must be consistent with some coherent set of probability assessments for the relevant uncertain outcomes, whether these assessments are made explicitly or not. Of course, the chance that a set of assessments that are made implicitly are actually coherent is exceedingly remote.

On the other hand, it is true that the theory does not solve the problem of eliciting probabilities, any more than the fact that there is a price at which you would just be willing to buy a new house allows you to determine without further ado what that price is. There is much work to be done on the elicitation problem, although many useful ideas and techniques already exist (for some of them, see the discussion of "Tactics" in section 3 below).

2) If the purpose of the causality assessment is to guide particular decisions, it is not necessary to evaluate each component probability precisely. As we shall show in section 3, a Bayesian causality assessment requires the evaluation of the probabilities of several different propositions. Suppose an assessor determines that his probability for one of these propositions lies somewhere between two numbers, say 1/3 and 1/2, but he finds it very difficult to evaluate it more specifically. It is always possible to go through the entire causality assessment, plugging in 1/3 wherever the probability in question appears, and then repeat the analysis using 1/2 instead. If the overall causality assessment is not very different, then there is no need to evaluate this particular probability more precisely. If it is, and some contemplated decision might hinge on the difference, then further work is unavoidable. The point is that such "sensitivity analyses" can always be carried out, with respect to the probabilities that prove to be the most difficult to assess; if it turns out to make no real difference which number in a certain range is used, then there is no need to introduce what seems like an arbitrary precision. If it matters, some creative tinkering using Bayes' Theorem, the Law of Total Probability or some other device is necessary to solve the problem.

3) What is the alternative to quantifying uncertainty? It is hard to see how the evidence about different factors and from different sources can be weighed and merged in a reasonable and nonarbitrary way without using quantitative methods that start with the explicit measurement of uncertainty. Certainly, as we argued in [14], no qualitative methods yet developed have come close to achieving this goal.

2. Collecting the Facts

The first step in the Bayesian approach is to collect the facts of the case. The assessor is required to identify the patient's clinical condition preceding the onset of the adverse event, the type of adverse event, the possible causes of the event, and the details of the evolution of the event:

1) Identify the clinical condition M and the type of the adverse event E_t.

M is a "generic" specification of the disease for which the patient is undergoing treatment, along with known co-morbidity (for example, M might denote pneumococcal pneumonia or stage four congestive heart failure). E_t is the "generic" type of the adverse event (for example, gastric ulcer or aplastic anemia), not a detailed description of the particular case at hand, which is denoted by E. Clearly, there is quite a bit of latitude in how specific these identifications can be; the important point is that they should be made explicitly at the beginning of the assessment, and thereafter whenever the assessment refers to "M" or "E_t" the meaning should remain the same. (Some guidelines for specifying M and E_t are presented in section 3 below.)

2) List the possible causes for the adverse event E.

The list should include possible drug causes (D_1 to D_n), the clinical condition M, nondrug treatment modalities, environmental exposures, and, of course, the possibility of some "unknown" cause.

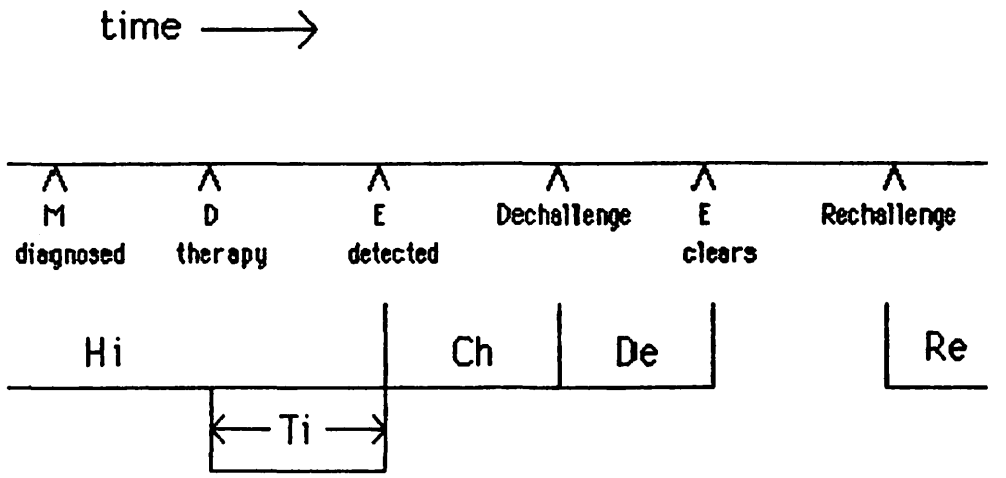
It is necessary to clarify what it means to say that something "causes E". The causes of primary interest in the causality assessment problem are drug causes: the proposition "D -> E" is true if E would not have happened as and when it did had D not been administered; this does not preclude the possibility that some attributes of the clinical condition M were also necessary for E to occur. The Bayesian approach requires that the alternative etiologies be listed in such a way that the elements of the list are mutually exclusive and exhaustive. Thus, if the listed possible causes are the drug D, the clinical condition M and "unknown", then the proposition "M -> E" means not only that E was a sequela to M, but also that E would have happened even if D had not been administered. That is, D-causation includes the possibility of an "interaction" between D and M, while M-causation expressly rules out D involvement.

Suspected drug interactions must be specifically incorporated into the list of possible causes. More precisely, when there is more than one drug causal candidate, if an interaction between drugs D_1 and D_2 is considered a priori as a possible cause, the list must include the hypotheses (D_1 alone), (D_2 alone) and (D_1 and D_2 jointly).

The list of causes is important in the Bayesian approach, since the approach works by partitioning the total probability, 1, among the various causal hypotheses, given all the elicited evidence. Thus, if a new etiological candidate is introduced to the list of causes, the causality assessment can change.

3) Record the relevant details about the case at hand.

It is most convenient to think about the case information that needs to recorded in terms of the chronology or time-course of the event E, an example of which can be illustrated as follows:



The symbols Hi, Ti, Ch, De and Re, which refer to different chronological classes of case information, are defined as follows:

Hi (historical information) contains information about the current patient that antedates the clinical episode initiated by E. Typically, Hi might include data about previous experiences with the suspect (and related) drugs and special demographic, behavioral, clinical or genetic risk classes for events of type E_t to which the patient belongs.

Ti describes the time of onset of E in relation to the administration of the drugs the patient has received, including when available the time-course of prodromal events like subclinical findings and early clinical signs and symptoms.

Ch (characteristics of E) refers to the period between the times of onset and dechallenge (or E clears, if no dechallenge occurs). Ch might include data about drug levels in tissues or body fluids as well as other details in clinical presentation, laboratory results, pathological findings or time-course that allow E to be more precisely described or classified .

De and Re refer to events in time periods initiated by dechallenge and rechallenge with the suspect drugs, when these occur. With respect to withdrawal of a particular drug D, De typically includes whether the symptoms associated with E abate when D is withdrawn (or its

dosage reduced), and, if so, the time-course and clinical characteristics of this response. Similarly, Re typically records whether an event of type E reappears following rechallenge, as well as the time it takes for this to happen if it does and any characteristics of the new event that provide differential etiological information. If a second dechallenge occurs following a positive response to rechallenge, a new class of information, De_2, must be introduced; similarly, Re_2, De_3 and so on may be necessary. In each case, all information about events in the relevant time period that can help distinguish between the various etiological candidates should be included in the appropriate chronological class.

The Bayesian approach requires the assessor to list what he knows about the case information in each of these classes, in response to the prompts given in Table 1 below (the questions in Table 1 are posed with respect to a particular suspect drug D; if more than one drug is a possible cause of D, repeat the questions with respect to each of them[3]). It is important to realize that the quantity of information in each of these five classes can vary widely from case to case. In particular, for most cases Hi and Ti contain important and sometimes abundant data. On the other hand, many events are irreversible, and so dechallenge and rechallenge cannot occur. Even if E is reversible, it may be sufficiently serious that rechallenge is not ethically feasible and so does not take place.

TABLE 1 GOES HERE

3. Evaluating the Evidence

After the facts have been collected, their evidentiary significance must be assessed. In this section we describe the goal of this assessment and the Bayesian strategy for achieving this goal. We also consider some tactics for implementing the Bayesian strategy.

The Goal

According to the Bayesian approach, the goal of causality assessment can be defined in the following way: for a drug D suspected as a cause of the adverse event E, calculate the posterior odds in favor of D-causation,

$$(1) \qquad \frac{P(D\rightarrow E \mid B, C)}{1 - P(D\rightarrow E \mid B,C)} \quad \text{-} \quad \frac{P(D\rightarrow E \mid B,C)}{P(D\not\rightarrow E \mid B,C)} .$$

Here, "D/>E" represents the proposition that the drug D did not cause the event E; that is, that E would have occurred as and when it did even if D had not been administered. C is the case

[3] See appendix 2 for a general technique for dealing with cases that have more than one possible drug cause, by considering a series of problems, each of which has only a single drug in its cause list.

information (that is, Hi, Ti, Ch, De, and Re) described in the previous section, and B is background information. B contains the fact that in the case at hand an event of type E_t has occurred in a patient with clinical condition M at some time after the drugs on the list of possible causes have been administered in a specified way. Unlike the case-specific information in C, the information in B makes no further reference to the particular patient whose case is currently undergoing assessment. In addition, B contains all the background information that the assessor might bring to bear to analyze any such case, including information about the drugs, their pharmacology and kinetics, indications, and risk factors that alter the chance of adverse effects.

The Strategy

Using de Finetti's measure, one could evaluate directly the probabilities in the numerator and denominator of the posterior odds displayed in equation (1), but, given the complexity of most causality assessment problems, such an act of global introspection could almost never be carried out consistently with all the opinions one holds about the meaning and relevance of the information in B and C. Thus, an alternative approach to the evaluation of the posterior odds is required. The strategy we adopt is to use the rules of Bayesian probability theory to decompose the posterior odds into a series of component factors, each of which require probability evaluations for propositions much more specific and accessible to the experience and knowledge of the evaluator than the proposition "D->E", and which are in principle subject to "predictive validity" tests as described in section 1 above. Once these component evaluations are carried out, there is only one coherent way to combine them to calculate the posterior odds in favor of D-causation, and that way is exhibited in equation (7) below.

We do not mean to imply that each of the component probability evaluations are "automatic". They may require a lot of thought, and the assessor may be far from confident in his answers. Nonetheless, we believe that these evaluations are addressable problems, and future work should lead to better techniques for their solution.

Strategy Step 1: Reduce to a Single Suspect Drug D

In many cases, the list of possible causes will include more than one drug or drug interaction. The first step in the Bayesian approach is to restrict to the case in which there is only one suspect drug, which we shall denote by D. This restriction involves no loss of generality, as we show in appendix 2 at the end of this paper (the reason is that Bayes' Theorem can be used to coherently merge the solutions of the causality assessment problems that arise when each suspect drug is treated in turn as though it were the only possible cause, into a solution to the overall assessment problem).

Strategy Step 2: Coherently Decompose the Posterior Odds

We now turn to the Bayesian decomposition of the posterior odds. The first part of this step is to apply Bayes' Theorem in odds form, as presented in section 1:

$$(2)\quad \frac{P(D \rightarrow E \mid B, C)}{P(D \not\rightarrow E \mid B, C)} = \frac{P(D \rightarrow E \mid B)}{P(D \not\rightarrow E \mid B)} \times \frac{P(C \mid D \rightarrow E, B)}{P(C \mid D \not\rightarrow E, B)}$$

posterior odds | prior odds | likelihood ratio

The first term on the right-hand side of equation (2), the prior odds, gives the odds in favor of drug causation taking into account just background information and disregarding any details about this particular patient and his adverse event. The second term on the right-hand side, the likelihood ratio, compares how likely are the details observed in this particular case under two competing etiological hypotheses: that the drug did and did not cause the adverse event E.

The advantage gained by the decomposition in equation (2) derives from the following consideration. As we have seen, the probabilities in the posterior odds refer to inherently unverifiable propositions: that the particular event E was or was not caused by D. On the other hand, the propositions whose probabilities appear in the prior odds and in the likelihood ratio are closely linked to predictive probabilities that can in principle be validated. We shall show how this applies to the likelihood ratio later; we now turn to the connection between the prior odds and predictive probabilities.

First, note that there is an important difference between the posterior and prior odds terms. Since both are evaluated conditionally on B, they both refer to a patient with clinical condition M who has been administered drug D in a particular way and who has at some unspecified time thereafter experienced an event of type E_t. However, the identity of the patient to whom the statements refer is different in the two terms. In the posterior odds, it is the particular patient for whom the causality assessment is being carried out, while in the prior odds it is a "generic" patient (for example, the "next" patient with M who suffers an event of this type after receiving D, for example) with the three defining properties -- clinical condition M, exposure to D and adverse event of type E_t. To reinforce this distinction, we shall substitute the expression "$D \rightarrow E_t$" for "$D \rightarrow E$" when the proposition refers to the "generic" patient with the defining characteristics rather than the particular patient at hand.

To see the connection between prior odds and predictive probabilities most clearly, it is easiest to focus on a special case. Imagine that records are available for a large group of patients with clinical condition M, and that the subgroup consisting of those patients who have received D is similar to the subgroup who have not, with respect to the distribution of any variables that are prognostic for the occurrence of events of type E_t (except, of course, for exposure to D).

The incidence of events of type E_t among those patients who take D is the sum of two components: the incidence of the events caused by D and the incidence of events not caused by D. Since the patients taking and not taking D are otherwise prognostically equivalent, the

incidence of events not caused by D in the patients taking D is the same as the incidence of all the events of type E_t in the patients not taking D. Since the prior probability that an event is caused by D is just the ratio of the incidence of events caused by D to the overall incidence of events among patients who receive D, we can summarize this discussion by the following equation[4]:

$$(3) \qquad P(D\text{->}E_t \mid B) = \frac{P(E_t|D) - P(E_t|D^c)}{P(E_t|D)} = 1 - [P(E_t|D^c)/P(E_t|D)]$$

where $P(E_t|D)$ is the probability that a patient who received D will experience an event of type E_t, and $P(E_t|D^c)$ is the probability that a patient who did not receive D will experience an event of this type. Both these probabilities clearly generate predictions of the incidence of events of type E_t, among future patients with M who do and do not receive D.

Usually, of course, the patients who do and do not receive D are not otherwise prognostically equivalent with respect to events of type E_t. Although more work is required to interpret $P(D\text{->}E_t|B)$ in terms of predictive probabilities in these cases, the principle that relates this probability to future incidences remains in force.

Now consider the probabilities in the numerator and denominator of the likelihood ratio. These call for an assessment of the probability that particular case details will occur, given that the event is and is not caused by the drug D. These probabilities are clearly about future observables (for the details of the "next" event of type E_t following administration of D to a patient with M), but since they are evaluated conditionally on an unobservable cause, they do not directly generate "validatable" predictions. But that problem is easily remedied. The Law of Total Probability can be used to combine the probabilities appearing in the likelihood ratio with the probabilities appearing in the prior odds, as follows:

$$(4) \quad P(C|B) = (P(C|B,D\text{->}E) \times P(D\text{->}E|B)) + (P(C|B,D/\text{>}E) \times P(D/\text{>}E|B)).$$

(Here, C refers to case details identical to the ones observed for the particular patient who experienced the event E.)

The probability on the left of equation (4) is a predictive probability, expressing the probability that the "next" patient with M who is administered D and experiences an event of type E_t will manifest case details C, while the four probabilities on the right are the components of the prior odds and the likelihood ratio. Since the prior odds components, P(D-

---

[4] In odds form,

$$P(D\text{->}E_t|B)/P(D/\text{>}E_t|B) = [P(E_t|D) - P(E_t|D^c)] \,/\, P(E_t|D^c).$$

›E|B) and P(D/›E|B), are assessable in terms of predictive probabilities and thus independently "validatable", equation (4) provides a predictive check for the likelihood ratio component probabilities.

We now turn to a further decomposition of the likelihood ratio. Trying to evaluate the probability of all the details in C simultaneously is too hard, since there is too much to think about at once. Therefore, repeatedly using the multiplicative condition for conditional probabilities given in section 1, we shall decompose the likelihood ratio into a series of factors each involving probabilities for propositions about only one of the five chronological classes Hi, Ti, Ch, De and Re:

$$(5)\qquad \frac{P(C \mid D \rightarrow E, B)}{P(C \mid D \not\rightarrow E, B)} = LR(Hi) \times LR(Ti) \times LR(Ch) \times LR(De) \times LR(Re)$$

Here, for example,

$$(6)\qquad LR(Hi) = \frac{P(Hi \mid D \rightarrow E, B)}{P(Hi \mid D \not\rightarrow E, B)};$$

$$LR(Ti) = \frac{P(Ti \mid D \rightarrow E, B, Hi)}{P(Ti \mid D \not\rightarrow E, B, Hi)} \quad \text{and}$$

$$LR(De) = \frac{P(De \mid D \rightarrow E, B, Hi, Ti, Ch)}{P(De \mid D \not\rightarrow E, B, Hi, Ti, Ch)}.$$

In words, LR(Hi) might be called the likelihood ratio factor evaluating historical information, and LR(De) the likelihood ratio factor evaluating dechallenge information, and so forth. The order in which these factors appear in equation (5) (and which factors appear as conditioning sets) is determined by chronology.

Putting equations (2) and (5) together gives the Bayesian decomposition of the posterior odds into the prior odds and five likelihood ratio factors:

$$(7)\qquad \frac{P(D\rightarrow E|B,C)}{P(D\not\rightarrow E|B,C)} = \frac{P(D\rightarrow E|B)}{P(D\not\rightarrow E|B)} \times LR(Hi) \times LR(Ti) \times LR(Ch) \times LR(De) \times LR(Re)$$

Tactics: Techniques for Implementing the Bayesian Strategy

In this section, we address the key tactical question for the Bayesian approach: how can an assessor evaluate the prior odds and the likelihood ratio factors? We begin by presenting some guidelines for defining M and E_t that make some of the probability evaluations easier. Next, we discuss four useful general techniques for evaluating probabilities. Finally, we offer specific suggestions for evaluating each of the six terms on the right-hand side of equation (7).

Warning: some readers may find some of the ideas in this section rough going. In such cases, it is probably best to skip or skim the section at the first reading and go directly to the example in section 4.

Defining M and E_t

The Bayesian strategy requires that the patient's clinical condition M and the type of adverse event E_t be unambiguously defined and that the definitions then be consistently applied in every subsequent probability evaluation. The level of specificity of these definitions can make a difference in how easy it is to carry out probability evaluations in which they play a role.

In particular, it is usually a good idea to attach a definite time horizon to the definition of the event type (that is, the definition of the event type is modified to include the requirement that the event occur sometime within a fixed amount of time -- the time horizon -- after the first administration of D). The role of the time horizon in the analysis is to separate the event under consideration from possible future occurrences of the same type of event. It is particularly useful in assessing the prior odds and the distribution for time to onset of the event as a function of the cause of the event. As a rule of thumb, we usually take as the horizon for a relatively common event a period at least as long as a "reasonable" time period for the event to occur as an adverse reaction to the suspect drug D, while for an uncommon event the horizon might be much longer. For example, if the event is diarrhea (as in the example in section 4), an appropriate time horizon might be one or two weeks, while if the event is Stevens-Johnson syndrome the horizon might be one year. The time horizon chosen can facilitate the assessment, but it does not affect the evaluation of the posterior odds in favor of drug causation.[5]

The assessor faces a choice between putting certain details of the patient's clinical condition into the definiton of M and placing them in Hi, which records information about the

---

[5] More accurately, we should say it should not affect the evaluation, and would not, if the assessor were coherent. Changing the time horizon will change the values of the different components of the posterior odds, but the changes compensate (for example, shortening the time horizon typically increases the prior odds in favor of drug causation, but lowers the likelihood ratio for timing proportionately).

patient's history before onset of E. Usually, the more general is the definition of M, the easier it is to assess the prior odds, but then the more specific becomes the information in Hi, and so the harder it is to assess LR(Hi). A useful rule is to include everything in M that is likely to have a substantial and relatively well-documented differential diagnostic effect (between D and non-D causation). For example, in one case we have examined, the patient was about to undergo renal transplantation surgery when pubic lice were detected and Kwell shampoo (1% gamma benzene hexachloride) applied. Eighteen hours after surgery, the patient died of a sudden cardiac arrest. In assessing the probability of Kwell-causation of the cardiac arrest, the patient's renal failure should certainly be included in M, even though it was not related to the Kwell-therapy. On the other hand, in a case involving an atopic patient who suffered Stevens-Johnson syndrome after undergoing sulfonamide treatment, the patient's atopy should be included in Hi rather than in the definition of M, because little information is available on the class of patients who are atopic and experience bacterial infections.

General Techniques for Probability Evaluation

Here are four general techniques for evaluating probabilities that can be used to advantage in the causality assessment context:

Conditioning: Sometimes, evaluating the probability of a proposition can seem difficult because the assessor's thoughts about the proposition depend on which of several other propositions are true. For example, if an assessor wants to determine the probability that an event of type E_t will occur as an adverse reaction to a drug D within a day of receiving a specified dosage of D, he might find that his assessment of this probability depends on the mechanism of the reaction (whether the reaction is immunologic or dose-dependent, for example). Or again, in assessing the probability that an event of type E_t that is not caused by D will occur in a specified time period, the evaluator might want to consider separately each possible alternative cause for the event.

In such cases, the Law of Total Probability can frequently be applied. First, the various possibilities on which the evaluation depends must be listed in such a way that one and only of one of them can be true (for example, the mechanism for the reaction may be immunologic, dose-dependent or "other"; or the alternative, nondrug causes for the event might be a viral infection for which the patient is being treated, some other, nondiagnosed infection or another, "unknown" cause). Then, the assessor must evaluate the probability of the proposition in question, conditional on each listed possibility. Next, he must evaluate the probability for each of the possibilities he has conditioned upon; this evaluation involves inherently unobservable propositions, but in our experience assessors often have little trouble partitioning their belief among a set of mutually contradictory mechanistic theories (the trickiest part is to decide how much probability to assign the catch-all "other" or "unknown"). Finally, the assessor puts together these two sets of evaluations according to the formula given in section 1 as the Law of Total Probability.

An example of this procedure will be presented in section 4 below.

Analogy: As explained in section 1, probability is just a measure of the assessor's uncertainty. Thus, it is sometimes possible for an assessor to evaluate a probability for a proposition by thinking about some other proposition about which his uncertainty is comparable and whose probability is easier to appraise.

For example, suppose you believe that the pharmacological mechanisms by which two related drugs can cause a particular kind of adverse reaction are very similar. Then it may be reasonable to suppose that the timing distributions for events of this type as adverse reactions to the two drugs are similar (in particular, say, the probability that E occurs within one day after receiving D_1, given that D_1 caused E, would be nearly the same as the same probability with respect to D_2). But the assessor may have much more experience with one of the drugs than with the other, in which case he might be quite confident about his assessment of the timing distribution corresponding to the familiar drug, which he can then transfer (perhaps with minor modifications) to the less familiar one.

As another example, suppose an assessor needs to evaluate his probability that the next infant receiving a course of therapy with a new "cillin"-type antibiotic drug will develop diarrhea. He can base this evaluation on the knowledge that reported incidences of diarrhea following therapy with other drugs in this family range from about 5 to 25%, with a mode of about 10% [10,11]; and so, if he is unaware of any feature of the new drug that would distinguish it from others in its class with respect to its propensity to cause diarrhea, he should assess the required probability, by analogy, at about 1/10.

Frequencies: Sometimes, an assessor may have access to observed frequencies that are clearly relevant to a probability evaluation problem he is trying to solve. For example, he might want to evaluate his probability that the next infant receiving a specified course of amoxicillin therapy will develop diarrhea, and he notes that a study monitoring outpatients in a large pediatric teaching hospital reported 130 cases of diarrhea out of 1320 patients within two days of beginning amoxicillin therapy [10]. Should he necessarily evaluate his probability as 13/132?

In general, the answer to this question is no. There are two primary reasons for this. The first has to do with the similarity between the patients for whom the probability evaluation is relevant and the patients upon whom the observed frequencies are based. The class of patients to which the probability evaluation refers is precisely specified by the conditions that appear to the right of the "|" in the statement of the probability: for example, if the probability in question is $P(E_t|D,M)$, the probability that a patient with clinical condition M who receives D in a specified way will experience an event of type E_t, the class consists of all patients who have the clinical condition defined by M and receive the course of therapy denoted by D. Differences in such factors as age, sex, severity of M, comorbidity, or the dosage of D, however, may compromise the applicability of frequencies reported in the literature to the specific class of patients for which probabilities are being assessed.

The second reason that probabilities may differ from observed frequencies has to do with chance variation. Even if the patients about whom frequency information is available are

characterized precisely as required for the probability evaluation problem, the probability and the observed frequency need not coincide, since the observed frequency reflects to some extent the vagaries of chance, especially if the sample size is small.

Now if the frequency information is based on patients in the class defined by the probability evaluation problem, and these patients have no other special defining characteristics and their number is large, then any coherent evaluation of the probability must be very close to the observed frequency. Otherwise, adjustments have to be made. Correcting for sample size is easy; dealing with the difference between the classes to which probability evaluation problems and observed frequencies refer is not.

Informally, we suggest the following solution to the problem. Use observed frequencies, when available, to provide an "anchor" or initial solution to a probability evaluation problem. Then think about the ways in which the class to which the frequency information refers differs from the class relevant to the probability evaluation, and decide what direction these differences suggest for changing the inital solution. (For example, if the observed frequencies for diarrhea following amoxocillin therapy were based only on infants in day-care centers, among whom one expects to find an elevated incidence of gastroenteritis, the frequencies should be adjusted downward to apply to the general infant population.) Finally, adjust the probability evaluation in the appropriate direction, perhaps using the conditioning or analogy techniques if they apply.

Sometimes, it is possible to use the connection between probability and frequencies to help evaluate probabilities, even when relevant observed frequencies are not available, by the following psychological ploy, the device of imaginary results. Suppose that an assessor has had a great deal of clinical experience with a particular kind of adverse event, and, for example, he must evaluate the probability that an event of this type will occur within one day of beginning D-therapy, given that it occurs sometime in the month after the therapy begins. Such an assessor might find it useful to draw upon his experience by imagining a great number of patients in the relevant class who have an event a month after beginning D-therapy, and asking himself what proportion of those patients he thinks will experience the event in the first day. If he can answer this question, he should use this proportion as his answer to the required probability evaluation problem.

Models: A model is a formal and general approach to probability evaluation. Models can be viewed as a systematic application of the ideas of conditioning and analogy. Because they can be constructed in accordance with the rules of probability theory, they give a framework for the coherent merger of different kinds of relevant information. As an example of the kind of model that would be helpful in the causality assessment context, think of the time of onset of a dose-dependent adverse reaction to a drug D. This time cannot be predicted with certainty, but it depends in part on certain pharmacological properties of the drug, physiological aspects of the reaction and specific attributes of the patient. A model for time to onset would specify how the mean reaction time depends on a particular set of drug-event-patient parameters, and it would also specify the pattern of the residual variability (which is not determined by the specified parameters). If such a model were constructed, the causality assessor would only need to specify the values of the input parameters for the

particular case at hand, and he could then use the model to compute the probability that an event of type E_t caused by drug D would occur just when the event E undergoing assessment occurred (which is the numerator of the likelihood ratio for timing).

Such models would be of great benefit in implementing the Bayesian strategy, because they would reduce the number of probability evaluations that an assessor would have to perform to carry out any particular causality assessment, they would substantially reduce the subjectivity in each of the remaining evaluations, and they would permit general predictive tests that would substantiate the models' and hence the whole Bayesian procedure's validity. We do not yet have such models, but one of the great advantages of the Bayesian approach is that it makes clear what models need to be developed, and it allows the incorporation of such models into the causality assessment procedure as they are developed.

Evaluating the Prior Odds

Equation (3) provides the most useful way to begin thinking about the prior odds. According to that equation, the prior odds can be regarded as a function of two incidence probabilities, $P(E_t|D)$ and $P(E_t|D^c)$, the first giving the incidence of events of type E_t among patients with M who receive the specified course of D-therapy, and the second giving the same incidence for an otherwise similar group of patients who do not receive D. Usually, such incidences are not known precisely. However, the assessor can always use the following tactic, whose precise formulation and probabilistic justification are presented in appendix 3: first, he expresses his uncertainty about the "true" incidences in the form of a probability distribution[6] for these two quantities; then, he uses the appropriate midpoints of these distributions (formally, their means -- see appendix 3) as his probabilities $P(E_t|D)$ and $P(E_t|D^c)$; finally, he computes the prior odds as a function of these probabilities according to the formula in equation (3).[7]

It is frequently possible to employ this tactic in a more informal way, particularly when the assessor has access to reasonably extensive and relevant frequency information (as is often the case for the incidence of events of type E_t when D is not administered, because estimates for the incidence of such events in the general population can frequently be obtained from the medical literature). When such frequency estimates exist, they can be

---

[6] A probability distribution for the "true" incidence gives values for the assessor's probabilities P(a ≤ "true" incidence ≤ b), for all a and b between 0 and 1. A distribution is usually specified by means of a density function, a nonnegative function f defined on the interval (0,1), such that P(a ≤ "true" incidence ≤ b) is obtained as the area under the graph of f between a and b (in particular, the total area under the graph of f is 1).

[7] In some cases, it might be easier to think about the difference $P(E_t|D) - P(E_t|D^c)$, rather than $P(E_t|D)$. The difference represents the incidence of events of type E_t following administration of D that are caused by the drug. The tactic, suitably modified, can be applied to estimate the "true" difference and the "true" incidence of events when D is not administered and to calculate the prior odds according to equation (3) from these two quantities.

used to evaluate the relevant probabilities directly, without constructing distributions for the "true" incidences, as in the discussion above on the general relation between probabilities and frequencies. But the cautions issued there hold: the assessor may need to make adjustments to the observed frequencies, since he is interested in the incidence among patients with clinical condition M, not the general population. If patients with M are at greater or less than average risk for events of type E_t, the assessor needs to modify the general incidences accordingly. Also, if the use of the drug D is high in the general population, the population incidence of events of type E_t represent mixtures of the incidences with and without D, and some adjustment is necessary before the observed frequencies can be used to give estimates of $P(E_t|D^c)$ alone.

Another informal method that sometimes works when information about the "true" incidence is sketchy involves applying the analogy technique for evaluating probabilities. For example, the assessor may feel that the connection between the drug D and the event E of interest is in the same range as some other drug-event connections, whose incidence figures are reasonably well-estimated in the literature, and he can adjust these incidences figures to give his $P(E_t|D)$.

But suppose neither of the informal substitution methods discussed in the previous two paragraphs works, and the assessor feels quite vague about what the "true" incidence for events of type E_t really is. As suggested in the first paragraph of this section, he should then try to assess a distribution that describes his uncertainty about the relevant "true" incidence. For example, he might feel fairly sure that the "true" incidence for events of type E_t following administration of D is somewhere between, say, 1/1000 and 1/10,000, but he cannot discriminate any more finely than this. Assuming that his uncertainly is approximately uniform in the "order of magnitude" scale, the argument given in appendix 3 suggests that he should assess $P(E_t|D)$ as 1/2558 (this is the mean of a distribution that is uniform in the log, or order of magnitude, scale, between 1/1000 and 1/10,000). The point is that the fact that the assessor's information is quite diffuse does not preclude evaluating a prior odds that accurately reflects his uncertainty.[8]

[8] Of course, when information is very diffuse and the assessor's opinion is correspondingly vague, his prior odds can change substantially if he gets access to new data that allows the "true" incidence to be estimated much more sharply. This in no way implies that the kind of calculation described above is "wrong"; only that the value of new information can be high when little is known.

Evaluating the Likelihood Ratio Factors

To evaluate the likelihood ratio factors, it is necessary to assess the probability for all the differentially diagnostic information elicited in response to the questions summarized in Table 1, given the contradictory hypotheses that D did and did not cause the adverse event E.

There is a conceptual difference between LR(Hi) and the other factors. Each of the other factors involves thinking in the forward direction, from a cause to its observable effects, while the information in Hi occurred before the event E, and so it is more difficult to think about how likely it was to be the case, conditional on the cause of an event that occurred after it did. Thus, we discuss LR(Hi) first and then turn to a general consideration of the other likelihood ratio factors.

LR(Hi): The representation for LR(Hi) given in equation (6) requires the assessor to think in "reverse chronology" (conditioning on the cause of E to evaluate the probability of events occurring before E), which is difficult. Thus, it is easier to think about LR(Hi) in terms of the following alternative representation, which is an immediate consequence of the multiplicative condition for conditional probabilities presented in section 1:

$$(8) \qquad LR(Hi) = \frac{P(D \rightarrow E|Hi,B) \,/\, P(D \not\rightarrow E|Hi,B)}{P(D \rightarrow E|B) \,/\, P(D \not\rightarrow E|B)}$$

In words, LR(Hi) is just the ratio of the odds in favor of drug causation taking into account the information in Hi (and no other case information) to the prior odds in favor of drug causation (ignoring the information in Hi). Seen in this way, the information in Hi serves as an adjustment to the prior odds, based on additional information about the patient that predates the occurrence of E. In effect, the relevant "reference set" in which to place the patient shifts from the general set of patients with M who experience an event of type E_t after the specified course of D-therapy, to those who share the same relevant history as the particular patient whose case is the subject of the assessment.

The adjustments to the prior odds required to evaluate LR(Hi) are often quite subjective, because the information that must be taken into account is too specific to expect to find readily assimilable observed frequencies based on large numbers of cases in the literature. Thus, it is useful to realize that certain classes of historical information, which seem to affect the incidence of events of type E_t, can be disregarded. In particular, according to equation (8), a datum in Hi will make LR(Hi) differ from 1 only if it affects the incidence of events of type E_t differentially between drug and nondrug causes. That is, if, say, the patient had some special attribute that doubled his risk for events of type E_t no matter what the cause, and the possession of this attribute by the patient was the only information in Hi, LR(Hi) would be 1. Thus, any such attributes can be disregarded in calculating LR(Hi).

Other Likelihood Ratio Factors: The likelihood ratio factors for timing, characteristics, dechallenge and rechallenge can best be evaluated by determining, separately, their numerators and denominators as given in equation (6) (for dechallenge). In carrying out these evaluations, the techniques of conditioning and analogy will be frequently applied. In particular, the probabilities given drug causation will depend on the mechanism of the adverse reaction, and the probabilities given nondrug causation will typically depend on what the alternative etiologies are; in both cases, conditioning on the appropriate entities is required.

The calculations involved in evaluating these factors are relatively straightforward, compared to the prior odds and LR(Hi), and further discussion will be deferred to the example given in the next section. Here, we consider only one issue. Notice that the information in chronologically preceding categories is conditioned upon when the probability for information in succeeding categories is calculated, as required by the multiplicative condition for coherence. For example, when calculating the probability that sulfonamide-induced Stevens-Johnson begins, say three days after onset of therapy, it is necessary to condition on historical information, like the fact that the patient under consideration is atopic. This successive conditioning at first sight seems to introduce a great deal of complexity to the evaluations, but in fact this is not generally so: all the probability evaluations required for these likelihood ratio factors are conditional on the cause of the event E (D, or some other cause); and conditional on the cause, the sets of information in the various categories are often independent, as would surely be the case with the timing information and the fact that the patient is atopic in the Stevens-Johnson example. When this conditional independence does not obtain, of course, the relevant conditioning information must be taken into account for a valid probability evaluation. This point is amply illustrated in the next section.

4. An Example

In this section, we apply the Bayesian approach to a case of suspected amoxicillin-induced diarrhea. The analysis is not based on a careful review of the literature; rather, it represents the clinical consensus of the authors of the paper, only one of whom (M.K.) has special expertise in this area. Nonetheless, we believe that the analysis provides a good introduction to the Bayesian approach and that the conclusion we draw is both essentially correct and consistent with all our opinions relating to the problem. The next paper in this series will analyze several more cases from the Bayesian point of view, and these cases will be substantially more problematic than the one analyzed here.

THE CASE

B.L. is a 17-month-old male day-care center attendee who on December 10 developed signs and symptoms of an upper respiratory tract infection with rhinorrhea and cough, but without fever or gastrointestinal symptoms. On the third day of his illness, his temperature rose to 39.4ºC, he became irritable, and he began to pull at his ears. He was seen by his pediatrician on that day and was diagnosed as having bilateral otitis media.

Treatment was initiated with amoxicillin suspension in a dose of 125 mg t.i.d. On the following day (the fourth day of the illness), B.L. had three watery bowel movements. By the fifth day, he was afebrile; the diarrhea continued, but without exacerbation. His mother telephoned the pediatrician, who suggested continuing the medication and encouraging fluid intake. B.L. remained afebrile and became less irritable and more playful, but the diarrhea persisted. The amoxicillin was discontinued on the thirteenth day (nine days after the diarrhea began). The diarrhea resolved on the fifteenth day (two days after dechallenge) and did not recur.

ANALYSIS

1) Preliminary Identifications

The Clinical Condition M: M is the upper respiratory tract infection (presumably viral), which by the third day is accompanied by fever and bilateral otitis media.

The Adverse Event Type E_t: A bout of frequent, loose stools, which we shall hereafter refer to as diarrhea, with a time horizon of one week from initiation of D-therapy.

Possible Causes of E: (1) Amoxicillin (denoted D hereafter); (2) Late-occurring GI symptoms secondary to the original infection (that is, M); (3) Coincidental gastroenteritis.

2) Case Information (Refer to Table 1)

Hi: There is no information about the patient's previous experience with D or events of type E_t. There is one attribute of the patient that places him at special risk to diarrhea from cause (3): he is a day-care attendee.

Ti: E began on the first day after D-therapy was initiated.

Ch: The only relevant information in this category is the duration of E: the diarrhea persisted for ten days before dechallenge took place (from the fourth to the thirteenth day of illness).

De: The diarrhea resolved, two days after dechallenge.

Re: No rechallenge occurred.

3) Evaluation of the Prior Odds

To calculate the prior odds, we estimate the numerator and denominator of the ratio on the right hand side of equation (3), using observed frequencies obtained from a study monitoring antibiotic-associated gastrointestinal symptoms in pediatric outpatients in Montreal (B.L.'s home) (some results from this study, but not the raw data that we use, are presented in [10,11]). In addition, we use data from Maricopa County, Arizona, presented in [20,21].

The first quantity we need to estimate is $P(E_t|D) - P(E_t|D^c)$; this difference estimates the "true" incidence of D-caused cases of events of type E_t. In the Montreal study, about 10% of the more than 1,300 patients receiving amoxicillin suffered from diarrhea within a week of beginning therapy. How many of these were drug-caused? To answer this question, we would like to know what the incidence of diarrhea in the same period would be among patients with M if they were treated with a drug as effective as amoxicillin that could not induce diarrhea as a side effect. Of course, no such drug exists. However, the lowest incidence of diarrhea in the study followed trimethoprim/sulfamethoxazole therapy, and was of the order of 2.5% in the first week of therapy. Thus, we estimate that the incidence (per child) of amoxicillin-caused diarrhea in the first week of therapy is at least 0.1 - 0.025 = 0.075. However, this is probably an underestimate, since some of the diarrhea following trimethoprim/sulfamethoxazole may represent adverse reactions to this drug. Thus, we must adjust our estimate of the incidence in nondrug-induced diarrhea downward somewhat, and as a result increase our estimate of the incidence of amoxicillin-induced diarrhea.

To decide how much of an adjustment to make, we argued along different lines. A lower bound for the incidence of nondrug-induced diarrhea can be obtained by thinking about the spontaneous occurrence of diarrhea (as in the definition of E_t), in which no drug involvement is possible because the affected child was taking no drugs prior to the outbreak of the diarrhea. We assume that one- to two-year old children experience approximately one such episode per year on average (this estimate is based primarily on the data in [20], adjusted upward to reflect pediatric experience in Montreal), which is equivalent to an incidence (per child) of about 0.019 per week. This figure must be increased somewhat, since it does not condition on the children having a viral infection (M), which increases the probability of developing diarrhea. We have thus determined that our estimate of the incidence (per child) of nondrug-induced diarrhea among children with M, in the week following initiation of amoxicillin therapy, should be greater than 0.019 and less than 0.025, and we adopt the value of 0.02. Thus, our estimate of the incidence (per child) of drug-caused diarrhea among such children in this time period is 0.1 - 0.02 = 0.08: this is our assessment of $P(E_t|D) - P(E_t|D^c)$.

To complete our evaluation of the prior odds, we need to assess $P(E_t|D^c)$. Our estimate for the incidence (per child) of nondrug-induced diarrhea in the calculation above was 0.02 cases per week. However, this estimate averages over the whole year. Now the incidence is substantially higher in the winter; in fact (see [21] for example) we estimate that the incidence of nondrug-induced diarrhea in an average winter week is three times greater than

the overall average. Hence our estimate of the incidence (per child) of nondrug-induced cases of diarrhea in the week following initiation of amoxicillin therapy is 0.06, which is our assessment for $P(E_t|D^c)$.

Thus, applying equation (3), we evaluate the prior odds in favor of D-causation as

$$\text{Prior Odds} = [P(E_t|D) - P(E_t|D^c)] / [P(E_t|D^c)] = 0.08/0.06 = 1.33$$

4) Evaluation of the Likelihood Ratio Factors

LR(Hi): We assume that day-care center attendees are 1.4 times as likely to suffer from non-drug-induced diarrhea (because of the greater exposure rate) as the general pediatric population (this estimate is primarily based on data in [21]), while they are at no added risk for drug-induced diarrhea. With these assumptions, the ratio as expressed in equation (8) is equal to 0.64. (To see this, go through the calculation of the prior odds again, with the only change being an increase in $P(E_t|D^c)$ by a factor of 1.4; note that this change affects both the numerator and denominator of the expression used to evaluate the prior odds. According to equation (8), LR(Hi) is the ratio of the result of this calculation to the prior odds as calculated in the previous paragraph.)

$$LR(Hi) = .64$$

LR(Ti): Recall that we condition on (as part of B) the information that the event E begins after D-therapy is initiated. Taking this into account, the pediatrician in our group (MK) assessed the following distributions for time to onset of an event of type E_t starting from the beginning of D-therapy, in a patient with M, given that an event of this type does occur (and hence, because of the time horizon, occurs within one week of the beginning of D-therapy):

Day of onset (i)	P(day i|D->E)	P(day i|M->E)[9]	P(day i|E coincidental)
1	.33	.33	.14
2	.33	.22	.14
3	.20	.15	.14
4	.07	.11	.14
5	.04	.09	.14
6	.02	.06	.14
7	.01	.04	.14

[9] This distribution was induced from the following distribution, which represents the probability of getting diarrhea beginning with the first day of M: .25 first day, .25 second day, .15 third day, .1 fourth day, .07 fifth day, .05 sixth day, .04 seventh day, .03 eighth day, .02 ninth day, .04 tenth day or later. The distribution given above is obtained from this one by conditioning on the diarrhea beginning between the third day (when D-therapy began) and the ninth day, a week later.

Here, "M->E" refers to cause (2) above, while "E coincidental" refers to cause (3). It should be noted that none of these timing distributions depend on the fact that B.L. attended a day-care center; that is, the information in Ti and the information in Hi are independent, given the cause of E.

Since E actually occurred on the second day of D-therapy, the numerator of the likelihood ratio factor for timing is .33. To evaluate the denominator, we need to evaluate two additional quantities, the probabilities for causes (2) and (3), given that D did not cause E. By the addition rule, these two numbers sum to 1. Because our estimate for the rate of coincidental diarrheas is 0.019 and our estimate for the rate of all nondrug-caused diarrheas is 0.02, we set P(M->E|D/>E) = 0.02 - 0.019 / 0.02 = .05 (so that the probability that M did not cause E, given that D did not cause E, is 0.95). Thus, the denominator of the likelihood ratio for timing, using the Law of Total Probability, is obtained as follows, where "day 2" is short for "onset of E occurs on day 2", "M↛E" represents the proposition that E was caused by a coincidental gastroenteritis, and every probability is conditional on background information B:

$$P(\text{day 2}|D/>E) = [P(\text{day 2}|M->E) \times P(M->E|D/>E)] + [P(\text{day 2}|M\not\to E) \times P(M\not\to E|D/>E)]$$
$$= (.22 \times .05) + (.14 \times .95)$$
$$= .144;$$

Thus,

$$\mathbf{LR(Ti) = P(\text{onset on day 2}|D->E) + P(\text{onset on day 2}|D/>E)}$$
$$= .33 + .144$$
$$= 2.3$$

Note that in the calculation for LR(Ti), the only part of the three timing distributions that was actually used was the probability they assigned to day 2, the day on which E actually occurred. Thus, if another assessor gives the same probabilities to day 2 but differs with us about the probabilites assigned to other days, the answer that assessor obtains for LR(Ti) for this case will agree with ours. The reason for assessing the entire distribution is that it provides a context for the one evaluation that counts, the probability assigned to what actually happened.

LR(Ch): To evaluate this factor, we need to calculate the probability that an event of type E_t will last at least ten days (corresponding to the fact that E lasted from the fourth day of the illness until the time of dechallenge, which occurred on the thirteenth day of the illness), given the hypotheses of drug and nondrug causation. Again, given the cause of E, the information whose probability we need to assess is independent of the information in Ti and Hi.

Assuming that D caused E, M.K., based on his clinical experience, assessed the probability that an event of type E_t would last at least ten days, given that the event was an adverse reaction to D as 0.7.

Assuming that D did not cause E, we believed that whether M caused E or E was coincidental to M, the time to resolution distribution was the same. This distribution was assessed by the pediatrician M.K. as follows: the probability that the diarrhea would end before five days is 0.3; between five and seven days, 0.4; on the eighth or ninth day, 0.15; on the tenth day, 0.05; on the eleventh day, 0.03; on the twelfth day, 0.02; and longer than twelve days, 0.05. Therefore, the probability that an event of type E_t would last at least ten days, given a nondrug cause, is 0.03 + 0.02 + 0.05 = 0.1.

$$\mathbf{LR(Ch) = 0.7 + 0.1 = 7}$$

LR(De): We need to calculate the probability that the diarrhea will resolve two days after dechallenge with D, given the hypotheses of drug and nondrug causation and given that the diarrhea had persisted until the time of dechallenge. These calculations are handled just as were the corresponding ones relative to the timing information. First, as explained in the discussion of Ch, we decided that there would be no difference in our distribution for the duration of the diarrhea whether it was caused by M or by a coincidental gastroenteritis. Using the distribution for duration of diarrhea above, conditioned to last at least 10 days, the probability that the diarrhea would resolve in the second day after dechallenge, given nondrug causation, is 0.02 / (0.03 + 0.02 + 0.05) = 0.2.

Assuming drug causation, the pediatrician M.K. assessed the following timing distribution for resolution: first day after dechallenge, 0.33; second day, 0.33; third day, 0.20; at least four days, 0.14 (this is just the distribution for time of onset, reversed).

$$\mathbf{LR(De) = 0.33 / 0.2 = 1.65}$$

LR(Re): Since no rechallenge occurred, this factor is equal to 1.

5) Conclusion: According to equation (7), to calculate the posterior odds in favor of D-causation, we must multiply together the prior odds and the likelihood ratio factors. Thus:

$$\begin{aligned}\textbf{Posterior odds} &= \textbf{Prior odds} \times \textbf{LR(Hi)} \times \textbf{LR(Ti)} \times \textbf{LR(Ch)} \times \textbf{LR(De)} \\ &= 1.33 \times 0.64 \times 2.3 \times 7 \times 1.65 \\ &= 22.6\end{aligned}$$

That is, the posterior odds are overwhelmingly in favor of drug causation (not surprisingly, since all the evidence, except that fact that the patient attended a day-care center, pointed in the same direction); the posterior probability of drug causation equals 22.6/23.6 or 0.96. Note that the strongest positive evidence was simply how long the diarrhea lasted before

dechallenge occurred; had the amoxicillin been immediately discontinued, the case for drug causation would have been far less convincing -- and had in addition the drug been trimethoprim/sulfamethoxazole instead of amoxicillin, so that the prior odds in favor of drug causation would be substantially lower (in fact, less than 1), the posterior odds would have favored a nondrug cause for the diarrhea.

5. Causality Assessment Criteria and the Bayesian Approach

In this section, we argue that the Bayesian approach to causality assessment satisfies, at least in principle, at least five of the six criteria introduced in [14]. The qualification "at least in principle" is necessary, because as presented in sections 2 and 3, the approach is difficult to implement and does not yet qualify as a standarized assessment method. Nonetheless, the Bayesian approach does substantially better with respect to these criteria, compared with other current causality assessment methods (see [14], where the schemes introduced in the following papers were evaluated according to the six criteria: [2,3,6,7,8,9,12,16,17,18]). Moreover, we believe that with further work it will be possible to develop techniques that make the approach easy to use without sacrificing what we see as its essential correctness.

We state each criterion (for justifications and discussion, see [14]) and then discuss how it relates to the Bayesian approach.

Criterion 1: Repeatability.

When the same "state of information" is used more than once as input a causality assessment method should produce the same "degree of belief" as output.

Discussion: In its present form, the Bayesian approach is just too complicated to achieve repeatability. The difficulty is not so much that there are too many probability assessments to make, each of which may vary from one evaluation session to the next even when the "state of information" does not change. Rather, the main problem is to determine in a standardized way what probabilities need to be evaluated for each of the components of the posterior odds. We believe that it will be possible to develop algorithmic methods based upon the Bayesian approach, in the context of specific drug-induced diseases (like cholestatic jaundice or Stevens-Johnson syndrome) or specific problem drugs (like digoxin). In such specific problem areas, "canonical" questions for eliciting the relevant information and operating with this information to determine the appropriate probabilities can be constructed. The more algorithmic the methods become, the more they will be able to satisfy this criterion.

Criterion 2: Explicitness

A causality assessment method should require that its user make explicit his "state of information", including the uncertainty he feels about each of its elements.

Discussion: The essence of the Bayesian approach is the explicit evaluation of uncertainty, so this requirement of the criterion is certainly satisfied. Since probabilities are most easily

evaluated the more exhaustively the problems are decomposed, the approach encourages the user to make explicit all his relevant information and the relations between its elements. Moreover, the techniques of conditioning and analogy provide a framework in which these relations can be systematically expressed.

Criterion 3: Explanatory Capability

A causality assessment method must "explain" how it reaches its conclusions; that is, it must make it clear to the user why it produced the output "degree of belief" from the information it elicited.

Discussion: The Bayesian approach reaches its conclusions from the component probability evaluations of the user by following the rules of prescriptive probability theory; that is, in the only coherent way. Moreover, the effect of the information in each factor on the output posterior odds is clear, since the posterior odds is just the product of the prior odds and the five likelihood ratio factors. In particular, it is easy to see at a glance which factors are the important ones in any assessment.

On the other hand, the method does not "explain" the probabilities that the user himself directly evaluates. Since these measure his own uncertainty about the propositions in question, in general he does not need an explanation for them. The exception occurs when the approach requires the user to evaluate probabilities about propositions that are meaningless to him. It should be possible to avoid this situation by creative decompositions and elicitation techniques, but we certainly cannot now guarantee to do so.

Criterion 4: Completeness

Any fact, theory or opinion that can affect an evaluator's belief that a drug D caused an adverse event E must be incorporable by a causality assessment method into the "state of information" on which the assessment is based.

Discussion: In principle, any such fact, theory or opinion can become the subject of a probability evaluation or the basis upon which such an evaluation is carried out. In particular, the Bayesian approach can deal with the three kinds of information singled out in [14] as essential, but not incorporable into other current assessment methods: uncertain information, quantitative information, and background information, especially epidemiological data about incidences and mechanistic theories from the basic sciences. Thus, the Bayesian approach already satisfies this criterion in principle; as better models are constructed to facilitate the incorporation of particular kinds of information, it will increasingly be able easily to satisfy it in practice.

Criterion 5: Etiological balancing

Methods cannot evaluate case data just in terms of their concordance or discordance with the hypothesis that the drug D caused the event E; rather, they must balance the likelihood of the data assuming that D caused E against the likelihood assuming alternative causes.

Discussion: This criterion is embedded in the architecture of the Bayesian approach.

Criterion 6: No a priori constraints on the effects of factors

A causality assessment method should not limit a priori the effect that information about any particular factor can have on the output "degree of belief".

Discussion: Each of the five likelihood ratio factors and the prior odds can range anywhere between zero and infinity, so the Bayesian approach places no a priori limits on their possible effects. Rather, these effects are only limited by the amount of information available and the state of the user's uncertainty about that information.

REFERENCES

1. Auriche, M. (1985). Approche Bayesienne de l'imputabilite des phenomenes indesirables aux medicaments, to appear in Therapie.

2. Blanc, S., P. Leunberger, J.P. Berger, E.M. Brooke and J.L. Schelling (1979). Judgements of trained observers on adverse drug reactions. Clinical Pharmacology and Therapeutics 25, 493-498.

3. Dangoumau, J., J.C. Evreux and J. Jouglard (1978). Methode d'imputabilite des effets adverses des medicaments. Therapie 33, 373-381.

4. deFinetti, B. (1972). Probability, Induction and Statistics (John Wiley, New York).

5. deFinetti, B. (1974), Theory of Probability (John Wiley, New York).

6. Emanueli, A. and G. Sacchetti (1980). An algorithm for the classification of untoward events in large scale clinical trials. Agents and Actions 7, 318-322.

7. Jones, J. (1982). Adverse drug reactions in the community health setting: Approaches to recognizing, counseling and reporting. Family and Community Health 5, 58-67.

8. Karch, F. and L. Lasagna (1977). Toward the operational identification of adverse drug reaction. Clinical Pharmacology and Therapeutics 21, 247-254.

9. Kramer, M.S., J.M. Leventhal, T.A. Hutchinson and A.R. Feinstein (1979). An algorithm for the operational assessment of adverse drug reactions: I. Background, description, and instructions for use. JAMA 242, 623-633.

10. Kramer, M.S., T.A. Hutchinson, L. Naimark, R. Contardi, K.M. Flegel and D.G. Leduc (1985). Antibiotic-Associated Gastrointestinal Symptoms in General Pediatric Outpatients. Pediatrics 76, 365-370.

11. Kramer, M.S., T.A. Hutchinson, L. Naimark, R. Contardi, K.M. Flegel and D.G. Leduc (1985). Adverse drug reactions in general pediatric outpatients. Journal of Pediatrics 106, 305-310.

12. Lagier, G., M. Vincens and A. Castot (1983). Imputabilite en pharmacovigilance: Principes de la methode appreciative ponderee et principales erreurs a eviter. Therapie 38, 303-318.

13. Lane, D. (1984). A probabilist's view of causality assessment. Drug Information Journal 18, 323-330.

14. Lane, D. and T. Hutchinson (1985). Assessing causality assessment methods. Submitted.

15. Lindley, D.V. (1971). Making Decisions (Wiley, New York).

16. Naranjo, C., U. Busto, E.M. Sellers, P. Sandor, I. Ruiz, E.A. Roberts, E. Janecek, C. Domecq and D.G. Greenblatt (1981). A method for estimating the probability of adverse drug reactions. Clinical Pharmacology and Therapeutics 30, 239-245.

17. Stephens, M. (1984). Assessment of causality in industrial settings. Drug Information Journal 18, 307-314.

18. Venulet, J., A. Ciucci and G.C. Berneker (1980). Standardized assessment of drug-adverse reaction associations -- rationale and experience. International Journal of Clinical Pharmacology, Therapy and Toxicology 18, 381-388.

19. Venulet, J., G.C. Berneker and A. Ciucci (eds) (1982). Assessing Causes of Adverse Drug Reactions (London, Academic Press).

20. Bartlett, A.V., M. Moore, G.W. Gary, K.M. Starko, J.J. Erben and B.A. Meredith (1985). Diarrheal illness among infants and toddlers in day care centers. I. Epidemiology and pathogens. Journal of Pediatrics 107, 495-502.

21. Bartlett, A.V., M. Moore, G.W. Gary, K.M. Starko, J.J. Erben and B.A. Meredith (1985). Diarrheal illness among infants and toddlers in day care centers. II. Comparison with day care homes and households. Journal of Pediatrics 107, 503-509.

TABLE 1: ELICITING CASE INFORMATION

The assessor should answer each of the following questions. If he is unsure of the correct answer, he should state the grounds and extent of his uncertainty (in probabilistic terms). While the answers to these questions will provide all of the relevant case information for many cases, any additional case information that can help differentiate between D and nondrug etiological candidates in a particular case under review should also be noted in the appropriate chronological period.

1. Hi:

   a. Has the patient taken D or similar drugs before? How frequently? On how many of these occasions did he experience an event of type E_t or another possible adverse reaction (describe, if different from E_t)?

   b. How frequently has the patient previously experienced events of type E_t without exposure to D or related drugs?

   c. Are there any attributes of the patient that place him at special risk to events of type E_t from any cause? If so, what are they, and from which causes is he at special risk?

2. Ti: When in relation to the course of D-therapy did the patient experience the event E (if available, give the time-course of all prodromal events)?

3. Ch:

   a. Is there any data about levels of D in tissues or body fluids during the time the patient experienced E? If so, what are they?

   b. Are there any distinctive details in clinical presentation, laboratory results, pathological findings or duration that can help differentially diagnose the cause of E? If so, what are they?

   c. Did the symptoms of E abate before dechallenge occurred? If so, how long after the time of onset of E?

4. De:

   a. Was D discontinued or its dosage reduced after the onset of E? If so, describe how and when.

   b. If dechallenge occurred, did the symptoms of E abate? If so, to what extent and when?

c. Were the symptoms of E treated directly? Was a specific antagonist to D administered? What was the result?

5. Re:

a. If dechallenge occurred, was the patient subsequently rechallenged with D? If so, when and in what dosage?

b. Did another event of type E_t (or related type) occur after rechallenge? If so, when?

APPENDICES

Appendix 1: Bayes' Theorem and the Law of Total Probability

Bayes' Theorem is derived from the multiplicative condition, which establishes the following equality:

(1) $P(A|B) \times P(C|A \text{ and } B) = P(C|B) \times P(A|C \text{ and } B)$.

Dividing both sides of (1) by P(C|B) gives

(2) $P(A|C \text{ and } B) = [P(C|A \text{ and } B) \times P(A|B)] / P(C|B)$,

which is Bayes' Theorem. To derive the odds form of Bayes' Theorem, apply the theorem with A^c in place of A to obtain

(3) $P(A^c|C \text{ and } B) = [P(C|A^c \text{ and } B) \times P(A^c|B)] / P(C|B)$.

Finally, divide equation (2) by equation (3), yielding

(4) $$\frac{P(A|C \text{ and } B)}{P(A^c|C \text{ and } B)} = \frac{P(C|A \text{ and } B)}{P(C|A^c \text{ and } B)} \times \frac{P(A|B)}{P(A^c|B)}.$$

The Law of Total Probability is derived as follows. First, since the propositions $A_1,\ldots,A_n$ are mutually contradictory and one of them is true,

(5) $C = (C \text{ and } A_1) \text{ or } (C \text{ and } A_2) \text{ or } \ldots \text{ or } (C \text{ and } A_n)$.

By the additivity condition, which obviously extends to n mutually contradictory propositions,

(6) $P(C|B) = P(C \text{ and } A_1|B) + \ldots + P(C \text{ and } A_n|B)$.

Now apply the multiplicative condition to each term on the right of (3):

(7) $P(C|B) = [P(C|B \text{ and } A_1) \times P(A_1|B)] + \ldots + [P(C|B \text{ and } A_n) \times P(A_n|B)]$.

Appendix 2: The One-Drug-at-a-Time Strategy

Let $D_1, \ldots, D_n$ represent all the (mutually contradictory) drug hypotheses on the list of causes, and group all the other hypotheses (for example: M, "other", etc.) as N (for nondrug).

Write $PO(D_i)$ for the posterior odds in favor of cause D_i. Now let A_i represent the hypothesis "D_i or N" (that is, the cause of E is either D_i, or a nondrug cause, and write $PO(D_i|A_i)$ for the posterior odds in favor of cause D_i, given A_i:

(1) $$PO(D_i|A_i) = P(D_i->E|B,C,A_i) / P(Di/>E|B,C,A_i).$$

Notice that $PO(D_i|A_i)$ solves the causality assessment problem for a case in which there is only one drug causal candidate, D_i.

Claim: The following formula gives the posterior odds in favor of cause D_i:

(2) $$PO(D_i) = PO(D_i|A_i) / [1 + \Sigma_{j=i} PO(D_j|A_j)].$$

That is, if the assessor calculates the conditional posterior odds in favor of cause D_j, $PO(D_j|A_j)$, for each possible drug cause D_j, then he can merge these condition odds to obtain the unconditional posterior odds in favor of cause D_i.

Proof of Claim: For succinctness, we omit B and C from the right of the "|" in all the probabilities that appear in this proof. By the multiplicative condition and the fact that A_i is equal to "D_i or N",

(3) $$P(D_i|A_i) = P(D_i \text{ and } A_i) / P(A_i)$$
$$= P(D_i) / [P(D_i) + P(N)].$$

Similarly,

(4) $$P(D_i^c|A_i) = P(N) / [P(D_i) + P(N)].$$

Thus,

(5) $$PO(D_i|A_i) = P(D_i) / P(N).$$

Applying equation (5) for each drug cause D_j, the right-hand side of equation (2) is equal to

(6) $$[P(D_i)/P(N)] / [1 + \Sigma_{j=i} P(D_j)/P(N)] = P(D_i) / [P(N) + \Sigma_{j=i} P(D_j)]$$
$$= P(D_i) / [1 - P(D_i)]$$
$$= PO(D_i),$$

which completes the proof of the claim.

Appendix 3: Predictive Probability and the "True" Incidence

The tactic discussed in section 3 under "Evaluating the Prior Odds" derives from the following theorem, due to deFinetti:

> Suppose $X_1, X_2, \ldots$ is a sequence of (0,1)-valued random variables, whose distribution is invariant under any reordering of the variables. Then
>
> (1) $\quad P(X_1 = 1) = \int_{[0,1]} y \, dG(y),$
>
> where G is the distribution function for the random variable
>
> $$Y = \lim_{n \to \infty} [(X_1 + \ldots + X_n) / n].$$
>
> (It is a conclusion of the theorem that this limit exists.)

The expression on the right of equation (1) is called the mean of the distribution G. If G has a density function g, then

$$\int_{[0,1]} y \, dG(y) = \int_{[0,1]} y \, g(y) \, dy.$$

To interpret the theorem in the context of this paper, $X_1, X_2, \ldots$ represents a sequence of future patients with clinical condition M who, say, are to receive a specified course of D-therapy and of whom nothing additional is known (that is, "generic" patients with the two stated properties. Because of their "genericness", the assessor's probability distribution for which of these patients will experience an event of type E_t does not depend on their labelling, so the theorem applies. Y represents the "true" incidence. Theus, the theorem implies that if the assessor evaluates his distribution G for the "true" incidence, his predictive probability that the next patient will experience an event of type E_t (that is, that $X_1=1$), or $P(E_t|D)$, is just the mean of the distribution G.